MATHEMATICS

# SOME INEQUALITIES FOR DISCRETE DISTRIBUTIONS WITH AN APPLICATION TO EIGENVALUES OF CERTAIN LINEAR OPERATORS

BY

M. H. VAN EMDEN

(Communicated by Prof. A. van Wijngaarden at the meeting of December 18, 1971)

## 1. Introduction

Consider a nonincreasing sequence $p$ of nonnegative numbers $p_1, p_2, \ldots$ of which $n$ are positive and whose sum is 1. The sum $\|p\|^2$ of the squares of these numbers ranges between $1/n$ in the case where $p_1 = p_2 = \ldots = p_n$ and 1 in the case where $p_1 = 1$. This suggests that, if we do not know the values of $p_1, \ldots, p_n$ but only of $\|p\|^2$, we may draw some conclusions about the numbers in the sequence. Thus, for large $\|p\|^2$, it seems obvious that $p_1$ cannot be small and that a more interesting lower bound may be found than the trivial one of $1/n$. On the other hand, if $\|p\|^2$ is small, this must also hold for $p_1$ which implies that $p_2$ cannot be very close to zero, and the same holds, perhaps only for even smaller $\|p\|^2$, for $p_3, p_4, \ldots$. Similar observations may be based on a known value of $|p| = p_1 \times p_2 \times \ldots$.

There is a theorem of Hardy, Littlewood, and Pólya that allows considerations like the above to be made quite precise: given $\|p\|^2$ it is possible to derive sharp upper and lower bounds for the sum of the first $i$ elements of the sequence $p$. Here, $i$ may be $1, 2, \ldots, n-1$, depending on the magnitude of $\|p\|^2$.

As an application we consider the eigenvalues of self-adjoint nonnegative definite operators in $n$-dimensional vector space. The eigenvalues are real, nonnegative and in certain cases their sum and the sum of their squares is easily computed. Therefore, the above-mentioned results for the elements of the sequence $p$ can be applied to the eigenvalues of such operators. The bounds will be shown to be the closest among those bounds that depend only on the sum and the sum of squares of the eigenvalues.

The theorem of Hardy, Littlewood, and Pólya applies not only to $\|p\|^2$ but, in general, to $\phi(p_1) + \ldots + \phi(p_n)$ where $\phi$ is a convex continuous function. Using $-\log$ as this function, and applying it in the finite-dimensional case, we find an upper bound for the sum of the first $i$ eigenvalues $(i = 1, \ldots, n-1)$ expressed in the trace and the determinant of the matrix of the operator.

## 2. An Ordering Among Distributions

We shall call a *discrete distribution* any sequence $p$ of nonnegative numbers $p_1, p_2, \ldots$ in nonincreasing order whose sum is 1. Let $n$ be the number of positive elements in $p$. We consider an ordering among distributions which we write $p \leqslant q$ ("$q$ majorizes $p$", where $q$ is a distribution), or equivalently $q \geqslant p$, which holds if and only if:

(1) $$p_1 + \ldots + p_i \leqslant q_1 + \ldots + q_i \text{ for } i = 1, 2, \ldots, n-1.$$

This is equivalent to

(2) $$p_{i+1} + \ldots + p_n \geqslant q_{i+1} + \ldots + q_n \text{ for } i = 1, \ldots, n-1.$$

We shall say that $b$ is a lower bound (an upper bound) of a set $S$ of distributions if and only if $b \leqslant p$ ($b \geqslant p$) for all $p$ in $S$.

Lemma 1. The set of distributions $p = p_1, \ldots, p_n$ ($n \geqslant 2$) for which $p_1 = a$ ($1/n < a \leqslant 1$) contains an upper bound, which is the distribution $q$: $q_1 = \ldots = q_i = a$, $q_{i+1} = 1 - ia$, $q_{i+2} = \ldots = q_n = 0$, where $i$ is the integer such that $ia \leqslant 1 < (i+1)a$.

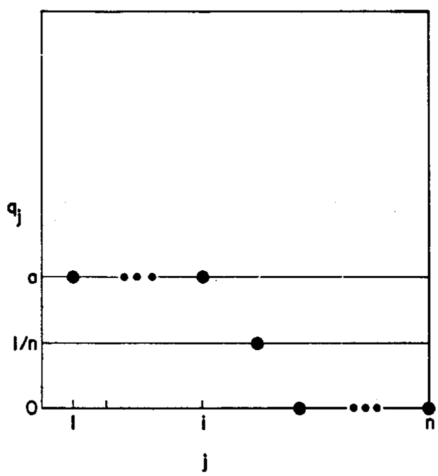


Fig. 1. An upper bound.

Proof: First, notice that $q$ belongs to the set; furthermore

$$p_1 + \ldots + p_j \leqslant jp_1 = ja = q_1 + \ldots + q_j \text{ for } j = 1, \ldots, i,$$

and

$$p_1 + \ldots + p_j \leqslant 1 = q_1 + \ldots + q_j \text{ for } j = i+1, \ldots, n,$$

which completes the proof by the definition (1) of the ordering.

Lemma 2. The set of distributions $p = p_1, \ldots, p_n$ ($n \geqslant 2$) for which $p_i = a$ ($0 \leqslant a \leqslant 1/i$, $i = 2, 3, \ldots$) contains an upper bound, which is the distribution $q$: $q_1 = 1 - (i-1)a$, $q_2 = \ldots = q_i = a$, $q_{i+1} = \ldots = q_n = 0$.

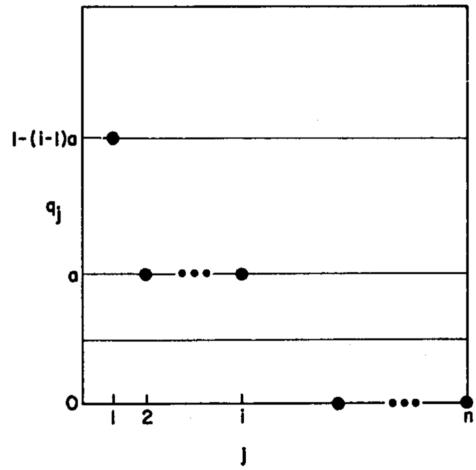


Fig. 2. An upper bound.

PROOF. First, notice that $q$ belongs to the set; furthermore

$$p_j+\ldots+p_n \geqslant 0=q_j+\ldots+q_n \text{ for } j=i+1, \ldots, n$$

and

$$p_j+\ldots+p_n \geqslant p_j+\ldots+p_i \geqslant (i-j+1)p_i=(i-j+1)a=q_j+\ldots+q_i= q_j+\ldots+q_n, \text{ for } j=2, \ldots, i,$$

which completes the proof by definition (2) of the ordering.

LEMMA 3. The set of distributions $p=p_1, \ldots, p_n$ $(n \geqslant 2)$ for which $p_1+\ldots+p_i=ia$ $(i/n<a \leqslant 1/a,\ i=1, \ldots, n-1)$ contains a lower bound which is the distribution $q$: $q_1=\ldots=q_i=a$, $q_{i+1}=\ldots=q_n=(1-ia)/(n-i)$.

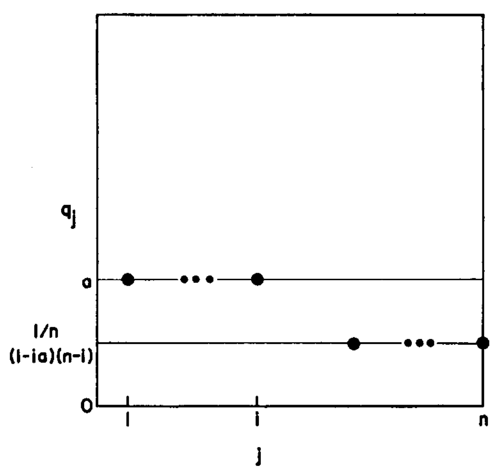


Fig. 3. A lower bound.

PROOF. First, notice that $q$ belongs to the set; furthermore

$$p_1+\ldots+p_j \geqslant ja = q_1+\ldots+q_j \text{ for } j=1, \ldots, i.$$

We also have

$$p_1+\ldots+p_j \geqslant q_1+\ldots+q_j \text{ for } j=i+1, \ldots, n$$

for suppose, on the contrary, that this inequality does not hold.

$$p_1+\ldots+p_j < q_1+\ldots+q_j \Rightarrow p_{j+1}+\ldots+p_n > q_{j+1}+\ldots+q_n =$$
$$=(n-j)(1-ia)/(n-i) \Rightarrow p_j > (1-ia)/(n-i) \Rightarrow$$
$$p_1+\ldots+p_j = ia + p_{i+1}+\ldots+p_j > ia+(j-i)(1-ia)/(n-i) = q_1+\ldots+q_j,$$

which is a contradiction. This completes the proof by the definition (1) of the ordering.

At this point the question presents itself whether the set of distributions for which $p_1+\ldots+p_i=ia$ $(1/n<a\leqslant 1/i,\ i=1, \ldots, n-1)$ contains an upper bound. The answer is negative. Notice that the set contains both the distribution $q$:

$$\text{(3)} \qquad \begin{cases} q_1=\ldots=q_{k-1}=a,\ q_k=1-(k-1)a, \\ q_{k+1}=\ldots=q_n=0, \end{cases}$$

where $k$ is the integer such that $(k-1)a\leqslant 1<ka$, and the distribution $r$:

$$\text{(4)} \qquad \begin{cases} r_1=1-(n-1)(1-ia)/(n-i), \\ r_2=\ldots=r_n=(1-ia)/(n-i). \end{cases}$$

An upper bound of the set would have to majorize both $q$ and $r$. Suppose, then, that $p\geqslant r$ and that $i$, $n$, and $a$ are such that $k<n$.

$p\geqslant r \Rightarrow p_1 \geqslant r_1 = 1-(n-1)(1-ia)/(n-i) \Rightarrow$ (because
$p_1+\ldots+p_i=ia$) $p_i \leqslant (1-ia)/(n-i) \Rightarrow$
$p_{i+1}+\ldots+p_k \leqslant (k-i)(1-ia)/(n-i) \Rightarrow$ (because $k<n$)
$p_1+\ldots+p_k<1 \Rightarrow p_1+\ldots+p_k<q_1+\ldots+q_k.$

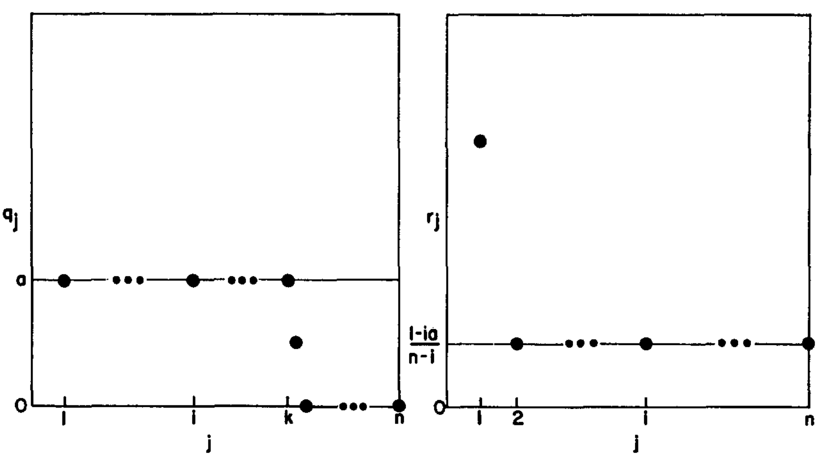


Fig. 4. Two distributions of a set whose upper bound is not in it.

This implies that $p$ does not majorize $q$ and that an upper bound of the set cannot belong to it. For this reason the set for which the next lemma states an upper bound is more restricted.

LEMMA 4. The set of distributions $p=p_1, \ldots, p_n$ $(n \geqslant 3)$ for which $p_1+\ldots+p_i=ia$ $(i=2, \ldots, n-1,\ 1/n<a \leqslant 1/i)$ and $p_i=b$ $((1-ia)/(n-i) \leqslant b \leqslant a)$ contains an upper bound which is the distribution $q$: $q_1=ia-(i-1)b$, $q_2=\ldots=q_{k-1}=b$, $q_k=1-ia-(k-i-1)b$, $q_{k+1}=\ldots=q_n=0$, where $k$ is the integer such that $(k-i-1)b \leqslant 1-ia<(k-i)b$.

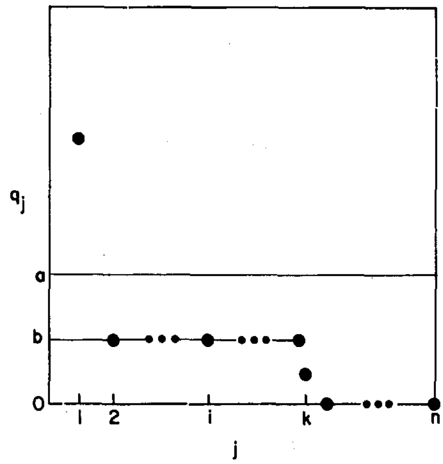


Fig. 5. An upper bound.

PROOF. First, notice that $q$ belongs to the set; furthermore

$$p_j+\ldots+p_n=p_j+\ldots+p_i+1-ia \geqslant (i-j+1)b+1-ia=q_j+\ldots+q_n$$

for $j=2, \ldots, i$;

$$p_j+\ldots+p_n=1-(p_1+\ldots+p_{j-1})=1-(ia+p_{i+1}+\ldots+p_{j-1}) \geqslant$$
$$\geqslant 1-ia-(j-i-1)b=q_j+\ldots+q_n \text{ for } j=i+1, \ldots, k;$$
$$p_j+\ldots+p_n \geqslant 0=q_j+\ldots+q_n \text{ for } j=k+1, \ldots, n,$$

which completes the proof according to the definition (2) of the ordering.

## 3. THE RESULT OF HARDY, LITTLEWOOD, AND PÓLYA

Apart from the relation (1), where $q$ majorizes $p$, Hardy, Littlewood, and Pólya considered a relation between distributions $p$ and $q$ (which they denoted by saying "$p$ is an average of $q$") defined to hold if and only if there are nonnegative numbers $a_{ij}$, $i=1, \ldots, n$, $j=1, \ldots, n$ such that

$$a_{i1}+\ldots+a_{in}=1,$$
$$a_{1i}+\ldots+a_{ni}=1,$$
$$p_i=a_{i1}q_1+\ldots+a_{in}q_n \text{ for } i=1, \ldots, n.$$

They proved (as Theorem 46 in [1]) that $q$ majorizes $p$ if and only if $p$ is an average of $q$. Using this result they proved (as Theorem 108 in [1]):

THEOREM 1.

$$\phi(p_1)+\ldots+\phi(p_n)\leqslant\phi(q_1)+\ldots+\phi(q_n)$$

holds for all convex continuous $\phi$ if and only if $q$ majorizes $p$.

## 4. Bounds Obtained From the Sum of Squares

Each of the previous lemmas combine with Theorem 1 to produce an inequality for elements of a distribution expressed in the sum of their squares. Each inequality is applicable to the eigenvalues of certain linear operators. The application is quite straightforward and is only explicitly given for Theorem 2 as a corollary to it.

THEOREM 2. Let $p=p_1, \ldots, p_n$ $(n\geqslant 2)$ be a distribution and $f(a)=ia^2+(1-ia)^2$ where, for each $a$ in $[1/n, 1]$, $i$ is the integer such that $ia\leqslant 1<(i+1)a$. The first element $p_1$ satisfies $f(p_1)\geqslant\|p\|^2$. This implies a lower bound $l_1$ for $p_1$ which is the value of $x$ in $[1/n, 1]$ satisfying $f(x)=\|p\|^2$. The lower bound is attained for $p=q$ as defined in Lemma 1, with $a=l_1$.

PROOF. According to Theorem 1 and Lemma 1 the maximum of $\|p\|^2$ over the set of distributions for which $p_1=a$ is $f(a)$ which is attained for $p=q$. Therefore $f(p_1)\geqslant\|p\|^2$ and $f(q_1)=\|p\|^2$. Because $f$ is a monotone nondecreasing function this implies the lower bound for $p_1$ mentioned in the theorem.

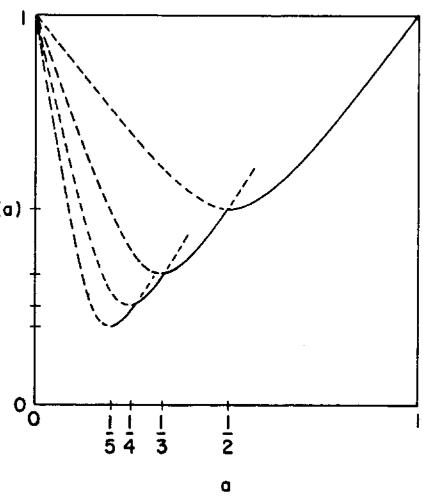


Fig. 6. Lower bound for first element.

Let $A$ be a self-adjoint nonnegative definite operator in either a finite $(n)$-dimensional vector space or in Hilbert space. In the latter case we assume $A$ to be also completely continuous. Let $y_1, \ldots, y_n$ be an orthonormal basis of the range of $A$ (which need not be finite). Then (see [3]) $tr(A)=(Ay_1, y_1)+\ldots+(Ay_n, y_n)$ exists (and we assume $A$ to be such that it is positive) and $A$ may be multiplied by a scalar $\alpha$ such that $tr(\alpha A)=1$. We may therefore assume without loss of generality that $tr(A)=1$.

COROLLARY. A largest eigenvalue $\lambda_1$ of $A$ satisfies $f(\lambda_1)\geqslant tr(A^2)$. This implies a lower bound $l_1$ for $\lambda_1$ which is the value of $x$ in $[1/n, 1]$ that satisfies $f(x)=tr(A^2)$. Any lower bound for $\lambda_1$ depending only on $tr(A^2)$ is at most $l_1$.

PROOF. The eigenvalues of $A$ are a countable set of nonnegative numbers. Call the positive among them $\lambda_1, \lambda_2, \ldots, \lambda_n$, in nonincreasing order. Then $tr(A)=\lambda_1+\ldots+\lambda_n$ and $tr(A^2)=\lambda_1^2+\ldots+\lambda_n^2$ (see [3] for these properties of $A$). By applying Theorem 2 to the eigenvalues of $A$ the bound mentioned in the corollary is obtained. The fact that there is a distribution, and therefore also an operator, for which the bound is attained implies that any bound, which depends on $tr(A^2)$ only, is at most $l_1$.

THEOREM 3. Let $p=p_1, \ldots, p_n$ $(n\geqslant 2)$ be a distribution and $f_i(a)=$ $=(i-1)a^2+(1-(i-1)a)^2$, $i=1, 2, \ldots, n-1$, $0\leqslant a<1/i$. The $i^{\text{th}}$ element $p_i$ satisfies $f_i(p_i)\geqslant \|p\|^2$. If $\|p\|^2>1/i$ this implies an upper bound $v_i$ for $p_i$, which is the smaller of the two values of $x$ satisfying $f_i(x)=\|p\|^2$. The upper bound is attained for $p=q$ as defined in Lemma 2, where $a=v_i$.

PROOF. According to Theorem 1 and Lemma 2 the maximum of $\|p\|^2$ over the set of distributions for which $p_i=a$ is $f_i(a)$, which is attained for $p=q$. Therefore $f_i(p_i)\geqslant\|p\|^2$. Because $f_i$ is a monotone decreasing function, this implies an upper bound for $p_i$.
For $i=2$ this theorem follows from Theorem 1. In some cases it gives results not obtained from Theorem 1. For instance, suppose $\|p\|^2=1/2$.

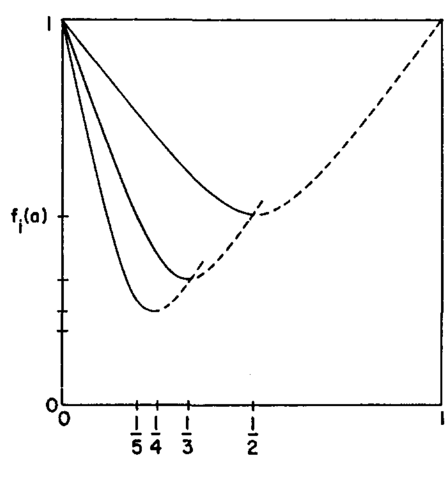


Fig. 7. Upper bounds for $p_2$, $p_3$ and $p_4$.

Theorem 1 gives: $\lambda_1 \geqslant 1/2 \Rightarrow \lambda_2 + \ldots + \lambda_n \leqslant 1/2 \Rightarrow \lambda_2 + \lambda_3 \leqslant 1/2 \Rightarrow \lambda_3 \leqslant 1/4$.
Theorem 2 gives $\lambda_3 \leqslant 1/6$.

THEOREM 4. Let $p = p_1, \ldots, p_n$ $(n \geqslant 2)$ be a distribution and $g_i(a) = ia^2 + (1-ia)^2/(n-i)$, $i = 1, \ldots, n-1$ and $1/n < a \leqslant 1$. The first $i$ elements of $p$ satisfy $g_i((p_1 + \ldots + p_i)/i) \leqslant \|p\|^2$. This implies an upper bound $u_i$ for their average, which is the larger of the two values of $x$ satisfying $g_i(x) = \|p\|^2$.
In the case of finite $n$ the upper bound is attained for $p = q$ as defined in Lemma 3, where $a = u_i$, for those values of $i$ for which $i \leqslant 1/\|p\|^2$. In the case of infinite $n$ the upper bound is not attained, but it is the least upper bound.

PROOF. According to Theorem 1 and Lemma 3 a lower bound of $\|p\|^2$ over the set of distributions for which $p_1 + \ldots + p_i = ia$ $(1/n < a \leqslant 1/i)$ is $g_i(a)$. This is a monotone increasing function and $g_i((p_1 + \ldots + p_i)/i) \leqslant \|p\|^2$; therefore implies an upper bound for $(p_1 + \ldots + p_i)/i$ if $\|p\|^2$ is known for those values $i \leqslant 1/\|p\|^2$.
In the case of finite $n$ the lower bound $g_i(a)$ for $\|p\|^2$ is attained and the bound for $(p_1 + \ldots + p_i)/i$ is attained for $p = q$, where $a = u_i$. In the case of infinite $n$, $g_i(a)$ is the greatest lower bound for $\|p\|^2$, and the bound for $(p_1 + \ldots + p_i)/i$ is the least upper bound.

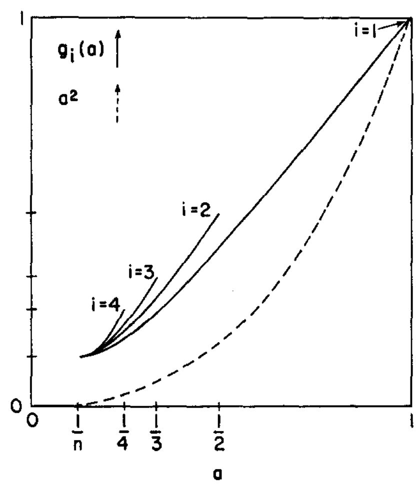


Fig. 8. Upper bounds for $p_1$, $(p_1+p_2)/2$, $(p_1+p_2+p_3)/3$ and $(p_1+p_2+p_3+p_4)/4$.

If $n = \infty$ and $i = 1$, the bound obtained from Theorem 4 is also obtained by observing that $p_1{}^2 \leqslant \|p\|^2$.

EXAMPLE. Let a completely continuous self-adjoint transformation $A$ in the Hilbert space of functions $f$ which are (Lebesgue) square-integrable on $[0, 1]$ be defined by

$$Af = 2 \int_0^1 \min(s, t) f(t) dt$$

and suppose that (from the context in which it arises) it is known that $A$ is positive definite. Call the eigenvalues $\lambda_1, \lambda_2, \ldots$ in decreasing order. Then (see [3]) the traces may be evaluated as follows:

$$\lambda_1 + \lambda_2 + \ldots = tr(A) = 2 \int_0 \min(t, t)dt = 1,$$

and

$$\lambda_1^2 + \lambda_2^2 + \ldots = tr(A^2) = 4 \int_0^1 \int_0^1 \min^2(s, t)dsdt = 2/3.$$

According to Theorem 4, the upper bound for $\lambda_1$ is $\sqrt{2/3} \leqslant .817$, which is the same bound as obtained from $\lambda_1^2 \leqslant \lambda_1^2 + \lambda_2^2 + \ldots = 2/3$. For matrices Theorem 4 gives a smaller upper bound; however, for increasing order the difference becomes smaller.

According to Theorem 2 the lower bound for $\lambda_1$ is $(1 + \sqrt{3}/3)/2 \geqslant .788$. Upper bounds for $\lambda_2, \lambda_3, \lambda_4$ and, for instance, $\lambda_{11}$ obtained by applying Theorem 3 are as follows:

$\lambda_2 \leqslant .212$ (from $\lambda_1 \geqslant .788$ we even have $\lambda_2 + \lambda_3 + \ldots \leqslant .212$),
$\lambda_3 \leqslant .099$,
$\lambda_4 \leqslant .064$

and

$\lambda_{11} \leqslant .019$.

THEOREM 5. Let $p = p_1, \ldots, p_n$ $(n \geqslant 2)$ be a distribution and $f(a) = (k-1)a^2 + (1-(k-1)a)^2$ where, for each $a$ in $[1/n, 1]$, $k$ is the integer such that $ka \geqslant 1 > (k-1)a$, and $h_i(a) = (n-1)(1-ia)^2/(n-i)^2 + (1-(n-1)(1-ia)/(n-i))^2$, $i = 1, \ldots, n-1$. The first $i$ elements of $p$ satisfy

$$\max\left(f((p_1 + \ldots + p_i)/i),\, h_i((p_1 + \ldots + p_i)/i)\right) \geqslant \|p\|^2.$$

This implies a lower bound for their average.

Let $x_1$ be the value of $x$ in $[1/n, 1]$ that satisfies $f(x) = \|p\|^2$ and let $x_2$ be the larger of the two values of $x$ that satisfy $h_i(x) = \|p\|^2$. If $x_1 < x_2$, then the bound is attained for $p = q$ (as defined in (3), where $a = x_1$). If $x_2 < x_1$ and if $n$ is finite, the bound is attained for $p = r$ (as defined in (4), where $a = x_2$); if $n$ is infinite, the bound is not attained but it is the greatest lower bound.

PROOF. According to Theorem 1 and Lemma 4 the maximum of $\|p\|^2$ over the set of distributions for which $p_1 + \ldots + p_i = ia$ $(i = 2, \ldots, n-1,$ $1/n < a \leqslant 1/i)$ and for which $p_i = b$ (where $(1-ia)/(n-i) \leqslant b \leqslant a$) is

$$(ia - (i-1)b)^2 + (k-2)b^2 + (1 - ia - (k-i-1)b)^2. \tag{5}$$

We need the maximum without the restriction to $p_i = b$ and we maximize (5) over the range allowed for $b$. It is a quadratic function in $b$ with a positive coefficient for $b^2$. Therefore, the maximum occurs either for

$b=(1-ia)/(n-i)$, and then (5) is equal to $h_i(a)$ or for $b=a$ and then (5) is equal to $f(a)$. Both are monotone nondecreasing functions of $a$, which implies the bound mentioned in the theorem.

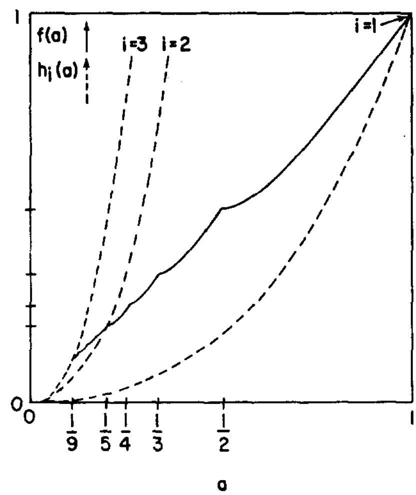


Fig. 9. Lower bounds for $p_1$, $(p_1+p_2)/2$ and $(p_1+p_2+p_3)/3$ in the case $n=\infty$.

For $n=\infty$ we have $h_i(a)=i^2a^2$ and $h_i(1/i^2)=1/i^2$. In this case it is easy to see whether $f$ or $h_i$ determines the bound: if $\|p\|^2 \leqslant 1/i^2$ then $f$ else $h_i$. Also, for any $n$, the lower bound is determined by $h$ if $\|p\|^2 \geqslant 1/i$.

## 5. Bounds Obtained From the Sum of Logarithms

If the convex continuous function $\phi$ in Theorem 1 is chosen to be $-\log$ then the following analog of Theorem 4 may be proved. In this theorem, the place of $\|p\|^2$ is taken by $-\log|p| = -\log(p_1) - \ldots - \log(p_n)$.

**Theorem 6.** Let $p=p_1, \ldots, p_n$ be a distribution and

$$d_i(a) = -i\log(a) - (n-i)\log((1-ia)/(n-i)).$$

The first $i$ elements of $p$ satisfy

$$d_i((p_1+\ldots+p_i)/i) \leqslant -\log|p|.$$

This implies an upper bound $w_i$ for their average, which is the larger of the two values of $x$ satisfying

$$d_i(x) = -\log|p|.$$

The upper bound is attained for $p=q$ as defined in Lemma 3 where $a=w_i$

Note that any value of $|p|$ gives an attainable upper bound for all $i=1, \ldots, n$; in Theorem 4 this is only the case for $1 \leqslant i \leqslant 1/\|p\|^2$.

By applying this theorem to the eigenvalues of a nonnegative definite Hermitian matrix $M$ of order $n$ with $tr(M)=1$ and eigenvalues $\lambda_1, \ldots, \lambda_n$ in nonincreasing order, we obtain the following

COROLLARY. The first $i$ eigenvalues of $M$ satisfy

$$d_i((\lambda_1+\ldots+\lambda_i)/i) \leqslant -\log |M|,$$

where $|M|$ is the determinant of $M$. This implies an upper bound $w_i$ for their average, which is the larger value of $x$ satisfying $d_i(x) = -\log |M|$. The upper bound is attained for any matrix having as eigenvalues the elements of $q$ in Lemma 3 (where $a=w_i$); any upper bound depending only on $|M|$ is at least $w_i$.

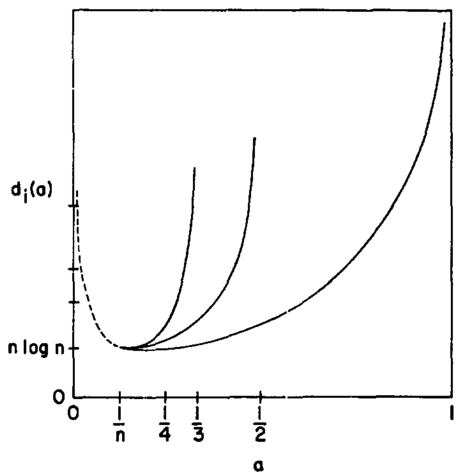


Fig. 10. Upper bounds for $p_1$, $(p_1+p_2)/2$ and $(p_1+p_2+p_3)/3$.

## 6. Concluding Remark

Much work has been done on bounds for eigenvalues of matrices and up till now we have not referred to any of it. One may well wonder whether the preceding theorems give results not obtainable by certain well-known theorems proved by methods more sophisticated than those employed in the present paper. For a survey of results on bounds for eigenvalues the reader is referred to Chapter 3 in HOUSEHOLDER [2] or Chapter 2 in WILKINSON [4].

As an example for comparison let us apply the Hoffman-Wielandt theorem (see [4]) to compare the diagonal elements of a symmetric positive definite matrix $C$ of order $n$ with its eigenvalues. Suppose that $tr(C)=1$. Let $A$ be the diagonal matrix of the diagonal elements of $C$ and let $B=C-A$. Call the eigenvalues of $A$, $B$, and $C$ respectively $\alpha_i$, $\beta_i$, and $\gamma_i$,

$i=1, \ldots, n$, in nonincreasing order. The Hoffman-Wielandt theorem asserts that, in this example,

$$(\gamma_1-\alpha_1)^2+\ldots+(\gamma_n-\alpha_n)^2 \leqslant \beta_1{}^2+\ldots+\beta_n{}^2. \tag{6}$$

The more nearly diagonal $C$ is (that is, the smaller $tr(B'B)=\beta_1{}^2+\ldots+\beta_n{}^2$ is) the closer are the bounds obtainable from the above inequality. Certainly, for a not too undiagonal $C$, any bounds obtained from our results are less interesting than those obtained by means of the Hoffman-Wielandt theorem.

But, how undiagonal can a matrix get? In the above example,

$$tr(A'A)+tr(B'B)=tr(C'C).$$

At one extreme, where $C$ is diagonal, $tr(B'B)=0$. At the other extreme, $tr(B'B)$ is large, and, for a given $tr(C'C)$, $tr(A'A)$ is small. The smallest possible value of $tr(A'A)=1/n$ (because $tr(A)=tr(C)=1$) and in that case $\alpha_i=1/n$, $i=1, \ldots, n$. In fact, any given $C$ is orthogonally equivalent to a matrix of which the diagonal elements are all equal to $1/n$. In that case we have equality in the inequality (6) because

$$\sum_{i=1}^{n}(\gamma_i-\alpha_i)^2=\sum_{i=1}^{n}\gamma_i{}^2-1/n=tr(C'C)-tr(A'A)=tr(B'B)=\sum_{i=1}^{n}\beta_i{}^2.$$

For this extreme case the inequality (6) gives no lower bound for $\lambda_1$, while Theorem 2 does give one. In this case it is still possible that $tr(C'C)$ is close to 1, so that the lower bound for $\lambda_1$ may even be quite close.

ACKNOWLEDGEMENT

I am indebted to Dr. A. J. HOFFMAN of the IBM T. J. WATSON Research Center for a helpful discussion on this subject.

*Mathematical Centre, Amsterdam,*
*The Netherlands IBM*
*T. J. Watson Research Center,*
*Yorktown Heights, New York*

## LITERATURE

1. HARDY, G. H., J. E. LITTLEWOOD and G. PÓLYA, Inequalities. Cambridge University Press (1934).
2. HOUSEHOLDER, A. S., The Theory of Matrices in Numerical Analysis. Blaisdell Publishing Company (1964).
3. NEUMANN, J. VON, Mathematical Foundations of Quantum Mechanics. Princeton University Press (1955).
4. WILKINSON, J. H., The Algebraic Eigenvalue Problem. Oxford University Press (1965).